Extremal words in the shift orbit closure of a morphic sequence

James D. Currie, Narad Rampersad, and Kalle Saari

Department of Mathematics and Statistics
University of Winnipeg
515 Portage Avenue
Winnipeg, MB R3B 2E9
Canada
j.currie@uwinnipeg.ca, {narad.rampersad, kasaar2}@gmail.com

Abstract. Given an infinite word $\mathbf{x}$ over an alphabet A, a letter $b \in A$, and a total order σ on A, we call the smallest word with respect to σ starting with b in the shift orbit closure $\mathcal{S}_{\mathbf{x}}$ of $\mathbf{x}$, if it exists, an *extremal word* of $\mathbf{x}$. In this paper we consider the extremal words of morphic words. If $\mathbf{x} = g(f^{\omega}(a))$ for some morphisms f and g, we give a simple condition on f and g that guarantees that all extremal words are morphic. An application of this condition shows that all extremal words of binary pure morphic words are morphic. Our technique also yields easy characterizations of extremal words of the Period-doubling and Chacon words and a new proof of the form of the lexicographically least word in the shift orbit closure of the Rudin-Shapiro word.

Keywords: Lexicographic order, morphic sequence, extremal word, Period-doubling word, Chacon word, Rudin-Shapiro word

1 Introduction

Given an infinite word $\mathbf{x} \in A^{\mathbb{N}}$, it is natural to inquire about the nature of the lexicographically extremal words in its shift orbit closure. We get different extremal words depending on the choice of the total order on the alphabet A and the initial letter of the extremal word. For example, if $A = \{\mathtt{0}, \mathtt{1}\}$ and $\mathbf{x}$ a Sturmian word, it is well-known that the extremal words with respect to $\mathtt{0} < \mathtt{1}$ are $\mathtt{0}\mathbf{c}$ and $\mathtt{10}\mathbf{c}$, where $\mathbf{c}$ is the characteristic word whose slope equals that of $\mathbf{x}$, and if we order $\mathtt{1} < \mathtt{0}$, then the extremal words are $\mathtt{1}\mathbf{c}$ and $\mathtt{01}\mathbf{c}$, see for example [14]. As another example, if $\mathbf{x}$ is k-automatic, then its extremal words are k-automatic as well [4]. For related results, see also [1,2,3,10]

The motivation for this paper comes from the following question: given a morphic sequence $\mathbf{x}$, when are the corresponding extremal words also morphic? While we are not able to solve the question in full generality, we give a fairly general condition (1) on the morphisms that generate $\mathbf{x}$ guaranteeing that the extremal words are morphic as well (Theorem 2). Using this condition we show

that the extremal words of all binary pure morphic words are morphic (Theorem 4). Then we move on to find characterizations of the extremal words of the Period-doubling (Theorem 5) and Chacon (Theorem 6) words. On our way to proving the main results, we show that if $\mathbf{x}$ is a pure morphic word generated by a morphism f and $\mathbf{t}$ is in the shift orbit closure of $\mathbf{x}$ such that $f(\mathbf{t}) = \mathbf{t}$, then $\mathbf{t}$ is morphic (Theorem 1).

2 Preliminaries

We will follow the standard terminology and notation of combinatorics on words as established, for example, in [5,13].

If A is an alphabet, then $A^{\mathbb{N}}$ denotes the set of all infinite words over A. If $X \subset A^*$, then X^ω denotes the set of all infinite words obtained by concatenating elements of X.

If $f\colon A^* \to A^*$ is a morphism such that $f(a) = ax$ for some letter $a \in A$ and a word $x \in A^+$ such that $f^n(x) \neq \varepsilon$, the empty word, for all $n \geq 0$, then there exists an infinite word $f^\omega(a) := \lim_{n\to\infty} f^n(a)$ such that $f^n(a)$ is a prefix of $f^\omega(a)$ for all $n \geq 0$, and it is called a *pure morphic word generated by* f. Notice that $f^\omega(a)$ is a fixed point of f, that is $f\bigl(f^\omega(a)\bigr) = f^\omega(a)$, but in general a fixed point of a morphism is not necessarily generated by the morphism (however, see Theorem 1).

If $c\colon A^* \to B^*$ is a coding, that is a letter-to-letter morphism, then $c(f^\omega(a))$ is called a *morphic sequence*. It is clear that all ultimately periodic sequences are morphic. The following result on morphic sequences is well-known, see Theorems 7.6.1 and 7.6.3 and Corollary 7.7.5 in [5].

Lemma 1. *Let $\mathbf{x} \in A^{\mathbb{N}}$ be a morphic sequence, $w \in A^*$, and $g\colon A^* \to B^*$ a non-erasing morphism. Then the words $w\mathbf{x}$, $w^{-1}\mathbf{x}$, and $g(\mathbf{x})$ are morphic.*

Let $\mathbf{x} \in A^{\mathbb{N}}$ be an infinite word. The set of factors of $\mathbf{x}$ is denoted by $F(\mathbf{x})$. We denote by $\mathcal{S}_\mathbf{x}$ the set of all infinite words $\mathbf{y} \in A^{\mathbb{N}}$ such that $F(\mathbf{y}) \subseteq F(\mathbf{x})$. Thus $\mathcal{S}_\mathbf{x}$ is the *shift orbit closure* of $\mathbf{x}$.

Now we are ready for the key definition of this paper. Let $f\colon A^* \to B^*$ be a morphism and $\mathbf{x} \in A^{\mathbb{N}}$. We will write

$$f \in \mathcal{M}_\mathbf{x} \tag{1}$$

if the following condition holds: for each letter $b \in A$, there exists a finite word $p_b \in B^+$ such that if $\mathbf{y} \in \mathcal{S}_\mathbf{x}$ begins with b, then $f(\mathbf{y})$ begins with p_b, and if $a \in A$ with $a \neq b$, then neither of p_a and p_b is a prefix of the other.

Example 1. Let us illustrate the above definition with a morphism appearing in [12]. Let f be given by $\mathtt{0} \mapsto \mathtt{02}$, $\mathtt{1} \mapsto \mathtt{02}$, and $\mathtt{2} \mapsto \mathtt{1}$, and let $\mathbf{x}$ be the unique fixed point of f. It is easy to see that if $\mathtt{0}\mathbf{y} \in \mathcal{S}_\mathbf{x}$, then $\mathbf{y}$ must begin with $\mathtt{2}$; hence $f(\mathtt{0}\mathbf{y})$ begins with $\mathtt{021}$. Similarly, if $\mathtt{1}\mathbf{y} \in \mathcal{S}_\mathbf{x}$, then $\mathbf{y}$ must begin with $\mathtt{0}$; hence $f(\mathtt{1}\mathbf{y})$ begins with $\mathtt{020}$. Finally, $f(\mathtt{2}\mathbf{y})$ begins with $\mathtt{1}$ regardless of $\mathbf{y}$. Therefore we may let $p_\mathtt{0} = \mathtt{021}$, $p_\mathtt{1} = \mathtt{020}$, and $p_\mathtt{2} = \mathtt{1}$, and consequently $f \in \mathcal{M}_\mathbf{x}$.

Let $\sigma = \sigma_A$ be a total order on an alphabet A, that is, a transitive and antisymmetric relation for which either (a, b) or (b, a) is in σ for all distinct letters $a, b \in A$. If $(a, b) \in \sigma$, we denote $a <_\sigma b$. The order σ extends to a lexicographic order on finite and infinite words over A in the usual way. Let $a \in A$ be a letter and $\mathbf{x} \in A^{\mathbb{N}}$ an infinite word in which a occurs. Then there exists a unique lexicographically smallest word in $\mathcal{S}_{\mathbf{x}}$ with respect to σ that begins with the letter a, and we will denote it by

$$\mathbf{l}_{a,\sigma,\mathbf{x}}.$$

Words of this form are collectively called the *extremal words* of $\mathbf{x}$ or $\mathcal{S}_{\mathbf{x}}$. We also denote by $\mathbf{s}_{a,\sigma,\mathbf{x}}$ the infinite word obtained from $\mathbf{l}_{a,\sigma,\mathbf{x}}$ by erasing the first letter, that is,

$$\mathbf{l}_{a,\sigma,\mathbf{x}} = a\,\mathbf{s}_{a,\sigma,\mathbf{x}}.$$

For the remainder of this section, let us fix a morphism $f \colon A^* \to A^*$. A word $u \in A^*$ is called *bounded under* f if there exists a constant $k > 0$ such that $|f^n(u)| < k$ for all $n \geq 0$. It is clear that every letter occurring in a bounded word is bounded. Let $B_f \subset A$ denote the set of bounded letters; the letters in $C_f := A \setminus B_f$ are said to be *growing under* f.

The following result is proved in [6, Prop. 4.7.62].

Lemma 2. *Suppose that $\mathbf{x} \in A^{\mathbb{N}}$ is a pure morphic sequence generated by f. There exists a finite subset Q of $C_f \times B_f^* \times B_f^* \times B_f^* \times B_f^* \times B_f^* \times C_f$ such that $F(\mathbf{x}) \cap C_f B_f^* C_f$ equals the set of all words of the form $c_1 y_1 z_1^k x z_2^k y_2 c_2$ with $(c_1, y_1, z_1, x, z_2, y_2, c_2) \in Q$ and $k \in \mathbb{N}$.*

Lemma 3. *Suppose that $\mathbf{x} \in A^{\mathbb{N}}$ is a pure morphic sequence generated by f. If $\mathbf{z} \in \mathcal{S}_{\mathbf{x}} \cap B_f^{\mathbb{N}}$, then $\mathbf{z}$ is ultimately periodic.*

Proof. Suppose that $\mathbf{z} \in \mathcal{S}_{\mathbf{x}} \cap B_f^{\mathbb{N}}$. If $\mathbf{x}$ has a suffix that is in $B_f^{\mathbb{N}}$, then it is ultimately periodic, which is proved in [6, Lemma 4.7.65], and then so is $\mathbf{z}$. Therefore we may assume that there are infinitely many occurrences of growing letters in $\mathbf{x}$. Let u_n be a sequence of factors of $\mathbf{x}$ such that u_n is a prefix of u_{n+1} for all $n \geq 1$ and $\mathbf{z} = \lim_{n\to\infty} u_n$. Since the first letter of $\mathbf{x}$ is necessarily growing and $\mathbf{x}$ has infinitely many occurrences of growing letters, it follows that each u_n is a factor of a word w_n such that $w_n \in C_f B_f^+ C_f \cap F(\mathbf{x})$. Since the set Q in Lemma 2 is finite, there exist letters $c_1, c_2 \in C_f$ and words $y_1, y_2, z_1, z_2, x \in B_f^*$ such that $w_{n_k} = c_1 y_1 z_1^{i_k} x z_2^{i_k} y_2 c_2$ for some subsequence n_k. By chopping off a prefix of length $|y_1|$ and a suffix of length $|y_2|$ from u_{n_k} if necessary, we may assume that each sufficiently long u_{n_k} is a factor of the biinfinite word $\mathbf{q} := {}^{\omega}z_1.xz_2^{\omega}$, where the the word x occurs in position 0. Now we have two possibilities: If there exists an integer $j \in \mathbb{Z}$ such that infinitely many u_{n_k} occurs in $\mathbf{q}$ in a position $\geq j$, then $\lim_{k\to\infty} u_{n_k}$ has suffix z_2^{ω}. If no such j exists, then $\lim_{k\to\infty} u_{n_k}$ has suffix z_1^{ω}. In the first case $\mathbf{z}$ has suffix z_2^{ω} and in the second case it has suffix z_1^{ω}.

Let $M_f \subset A$ denote the set of letters b such that $f^i(b) = \varepsilon$ for some integer $i \geq 1$, and let $t \geq 1$ be the smallest integer such that $f^t(b) = \varepsilon$ for all $b \in M_f$. Let

$$G_f = \{ f^t(a) \mid a \in A \text{ such that } f(a) = xay \text{ for some } x, y \in M_f^* \}$$

Notice that each word $f^t(a)$ in G_f is a finite fixed point of f because

$$f^t(a) = f^{t-1}(x) \cdots f(x)xayf(y) \cdots f^{t-1}(y).$$

In particular, all words in G_f are bounded. The following result is by Head and Lando [11], see also [5, Theorem 7.3.1].

Lemma 4. *Let* $\mathbf{t} \in A^{\mathbb{N}}$ *be an infinite word. We have* $f(\mathbf{t}) = \mathbf{t}$ *if and only if at least one of the following two conditions holds:*

(a) $\mathbf{t} \in G_f^{\omega}$*; or*
(b) $\mathbf{t} = wf^{t-1}(x) \cdots f(x)xayf(y)f^2(y) \cdots$ *for some* $w \in G_f^*$ *and* $a \in A$ *such that* $f(a) = xay$ *with* $x \in M_f^*$ *and* $y \notin M_f^*$*.*

Lemma 5. *Let* $f \colon A^* \to A^*$ *be a morphism. If* $\mathbf{t} \in A^{\mathbb{N}}$ *can be written in the form* $\mathbf{t} = wxf(x)f^2(x)f^3(x) \cdots$*, where* $w \in A^*$ *and* $x \notin M_f^*$*, then* $\mathbf{t}$ *is morphic.*

Proof. Let b be a new letter that does not occur in A. Then the infinite word $bxf(x)f^2(x) \cdots$ is morphic as it is generated by a morphism $g \colon (A \cup \{b\})^* \to (A \cup \{b\})^*$ for which $g(b) = bx$ and $g(a) = f(a)$ for all $a \in A$. Thus it follows from Lemma 1 that $\mathbf{t}$ is morphic.

Theorem 1. *Let* $f \colon A^* \to A^*$ *be a morphism, and suppose that* $\mathbf{x} \in A^{\mathbb{N}}$ *is a pure morphic word generated by* f*. If* $\mathbf{t} \in \mathcal{S}_{\mathbf{x}}$ *satisfies* $f(\mathbf{t}) = \mathbf{t}$*, then* $\mathbf{t}$ *is morphic.*

Proof. According to Lemma 4, either $\mathbf{t}$ is in G_f^{ω} or it is of the form

$$\mathbf{t} = wf^{t-1}(x) \cdots f(x)xayf(y)f^2(y) \cdots .$$

In the former case $\mathbf{t} \in B_f^{\mathbb{N}}$, so $\mathbf{t}$ is ultimately periodic by Lemma 3, and thus morphic. In the latter case $\mathbf{t}$ is morphic by Lemma 5.

3 Main theorem

Lemma 6. *Let* $\mathbf{x} \in A^{\mathbb{N}}$ *be an infinite word and* $f \colon A^* \to B^*$ *a morphism. If* $\mathbf{y} \in B^{\mathbb{N}}$ *is in* $\mathcal{S}_{f(\mathbf{x})}$*, then there exist a letter* $a \in A$ *and an infinite word* $\mathbf{z}$ *such that* $a\mathbf{z} \in \mathcal{S}_{\mathbf{x}}$ *and* $\mathbf{y} = uf(\mathbf{z})$*, where* u *is a nonempty suffix of* $f(a)$*.*

Proof. Let L_n denote the length-n prefix of $\mathbf{y}$; then L_n is a factor of $f(\mathbf{x})$ by the definition of $\mathcal{S}_{f(\mathbf{x})}$. Consequently, if $n \geq \max_{a \in A}|f(a)|$, there exist letters $a_n, b_n \in A$ and a word $v_n \in A^*$ such that $a_n v_n b_n$ occurs in $\mathbf{x}$ and we have $L_n = s_n f(v_n) p_n$, where s_n is a nonempty suffix of $f(a_n)$ and p_n is a possibly

empty prefix of $f(b_n)$. Since there are only finitely many different possibilities for a_n and s_n, there exists a letter $a \in A$ and a word u such that $a_{n_i} = a$ and $s_{n_i} = u$ for infinitely many n_i. The set of words $\{v_{n_i}\}$ being infinite, König's Lemma implies that there exists an infinite word $\mathbf{z}$ such that every prefix of $\mathbf{z}$ is a prefix of some v_{n_i}. Since each of av_{n_i} is a factor of $\mathbf{x}$, we have $a\mathbf{z} \in \mathcal{S}_{\mathbf{x}}$. Furthermore, since each prefix z of $\mathbf{z}$ is a prefix of some v_{n_i}, the word $uf(z)$ is a prefix of $\mathbf{y}$, and consequently $\mathbf{y} = uf(\mathbf{z})$.

Lemma 7. *Let $f\colon A^* \to B^*$ be a morphism and $\mathbf{x} \in A^{\mathbb{N}}$ such that $f \in \mathcal{M}_{\mathbf{x}}$. Let $b \in B$ be a letter that occurs in $f(\mathbf{x})$ and let ρ be a total order on B. Then there exist a total order σ on A, a letter $a \in A$, and a possibly empty proper suffix v of $f(a)$ such that*

$$\mathbf{s}_{b,\rho,f(\mathbf{x})} = vf(\mathbf{s}_{a,\sigma,\mathbf{x}}). \tag{2}$$

Proof. By Lemma 6, we can write $\mathbf{l}_{b,\rho,f(\mathbf{x})} = b\mathbf{s}_{b,\rho,f(\mathbf{x})} = uf(\mathbf{z})$, where u is a nonempty suffix of $f(a)$ for some $a \in A$ and $a\mathbf{z} \in \mathcal{S}_{\mathbf{x}}$. Since $f \in \mathcal{M}_{\mathbf{x}}$, there exist words $p_x \in B^+$ for every $x \in A$ such that $p_x \neq p_y$ whenever $x \neq y$. Thus we can define a total order σ on A such that, for all letters $x, y \in A$, we have $x <_\sigma y$ if and only if $p_x <_\rho p_y$.

We claim that $\mathbf{z} = \mathbf{s}_{a,\sigma,\mathbf{x}}$. If this is not the case, then $\mathbf{z} >_\sigma \mathbf{s}_{a,\sigma,\mathbf{x}}$ because both $a\mathbf{z}$ and $a\mathbf{s}_{a,\sigma,\mathbf{x}} = \mathbf{l}_{a,\sigma,\mathbf{x}}$ are in $\mathcal{S}_{\mathbf{x}}$ and $\mathbf{l}_{a,\sigma,\mathbf{x}}$ is the smallest word in $\mathcal{S}_{\mathbf{x}}$ starting with the letter a. Therefore $\mathbf{z} = wy\mathbf{t}$ and $\mathbf{s}_{a,\sigma,\mathbf{x}} = wx\mathbf{t}'$ with $x, y \in A$ satisfying $y >_\sigma x$. Since $f(x\mathbf{t})$ begins with p_x and $f(y\mathbf{t}')$ begins with p_y and neither of p_x and p_y is a prefix of the other, we have $f(y\mathbf{t}) >_\rho f(x\mathbf{t}')$ by the definition of σ, and this gives

$$\mathbf{l}_{b,\rho,f(\mathbf{x})} = uf(\mathbf{z}) = uf(w)f(y\mathbf{t}) >_\rho uf(w)f(x\mathbf{t}') = uf(\mathbf{s}_{a,\sigma,\mathbf{x}}).$$

But this contradicts the definition of $\mathbf{l}_{b,\rho,f(\mathbf{x})}$ because $uf(\mathbf{s}_{a,\sigma,\mathbf{x}})$ starts with the letter b and is in $\mathcal{S}_{f(\mathbf{x})}$. Therefore we have shown that $\mathbf{z} = \mathbf{s}_{a,\sigma,\mathbf{x}}$, and so (2) holds with $v = b^{-1}u$.

Lemma 8. *Let $f\colon A^* \to A^*$ be a morphism and $\mathbf{x} \in A^{\mathbb{N}}$ such that $f \in \mathcal{M}_{\mathbf{x}}$ and $f(\mathbf{x}) = \mathbf{x}$. Then for any total order ρ on A and a letter $b \in A$ occurring in $\mathbf{x}$, there exist a total order σ on A, a letter $a \in A$, words $u, v \in A^*$, and integers $k, m \geq 1$ such that*

$$\mathbf{s}_{b,\rho,\mathbf{x}} = uf^k(\mathbf{s}_{a,\sigma,\mathbf{x}}) \qquad \textit{and} \qquad \mathbf{s}_{a,\sigma,\mathbf{x}} = vf^m(\mathbf{s}_{a,\sigma,\mathbf{x}}). \tag{3}$$

Proof. Since $f(\mathbf{x}) = \mathbf{x}$, Lemma 7 implies that $\mathbf{s}_{b,\rho,\mathbf{x}} = v_0 f(\mathbf{s}_{a_1,\sigma_1,\mathbf{x}})$ for some total order σ_1 on A, a letter $a_1 \in A$, and a possibly empty suffix v_0 of $f(a_1)$. By applying Lemma 7 next on $\mathbf{s}_{a_1,\sigma_1,\mathbf{x}}$ and further, we get a sequence of identities

$$\mathbf{s}_{a_k,\sigma_k,\mathbf{x}} = v_k f(\mathbf{s}_{a_{k+1},\sigma_{k+1},\mathbf{x}}) \qquad (k \geq 0),$$

where we denote $a_0 = b$ and $\sigma_0 = \rho$. Therefore,

$$\mathbf{s}_{a_k,\sigma_k,\mathbf{x}} = v_k f(v_{k+1}) \cdots f^{m-1}(v_{k+m-1}) f^m(\mathbf{s}_{a_{k+m},\sigma_{k+m},\mathbf{x}}),$$

for all integers $k \geq 0$ and $m \geq 1$. Since there are only finitely many different letters and total orders on A, there is a choice of k and m such that $a_k = a_{k+m}$ and $\sigma_k = \sigma_{k+m}$. Thus by denoting $a = a_k$, $\sigma = \sigma_k$,

$$u = v_0 f(v_1) \cdots f^{k-1}(v_{k-1}), \quad \text{and} \quad v = v_k f(v_{k+1}) \cdots f^{m-1}(v_{k+m-1}),$$

we have the identities in (3).

Lemma 9. *Let $f \colon A^* \to A^*$ be a morphism and $\mathbf{x} \in A^{\mathbb{N}}$ such that $f \in \mathcal{M}_{\mathbf{x}}$ and $f(\mathbf{x}) = \mathbf{x}$. Then for any total order ρ on A and a letter $b \in A$ occurring in $\mathbf{x}$, there exist a finite word $w \in A^+$, an infinite word $\mathbf{t} \in \mathcal{S}_{\mathbf{x}}$, and an integer $m \geq 1$ such that*

$$\mathbf{l}_{b,\rho,\mathbf{x}} = w\mathbf{t} \tag{4}$$

and either

$$\mathbf{t} = f^m(\mathbf{t}) \tag{5}$$

or

$$\mathbf{t} = \lim_{n\to\infty} x f^m(x) f^{2m}(x) \cdots f^{nm}(x) \tag{6}$$

for some finite word $x \in A^+$.

Proof. According to Lemma 8, there exist a total order σ on A, a letter $a \in A$, words $u, v \in A^*$, and integers $k, m \geq 1$ such that

$$\mathbf{s}_{b,\rho,\mathbf{x}} = u f^k(\mathbf{s}_{a,\sigma,\mathbf{x}}) \qquad \text{and} \qquad \mathbf{s}_{a,\sigma,\mathbf{x}} = v f^m(\mathbf{s}_{a,\sigma,\mathbf{x}}).$$

Denote $w = bu$ and $\mathbf{t} = f^k(\mathbf{s}_{a,\sigma,\mathbf{x}})$. Then $\mathbf{t} \in \mathcal{S}_{\mathbf{x}}$ and Eq. (4) holds. By denoting $x = f^k(v)$, we get $\mathbf{t} = x f^m(\mathbf{t})$. If $x = \varepsilon$, then we have $\mathbf{t} = f^m(\mathbf{t})$, and Eq. (5) holds. If $x \neq \varepsilon$, then

$$\mathbf{t} = x f^m(\mathbf{t}) = x f^m(x) f^{2m}(\mathbf{t}) = \cdots = x f^m(x) f^{2m}(x) \cdots f^{nm}(x) f^{(n+1)m}(\mathbf{t}),$$

for all integers $n \geq 0$. The morphism f is nonerasing because $f \in \mathcal{M}_{\mathbf{x}}$, and therefore the words $x f^m(x) f^{2m}(x) \cdots f^{nm}(x)$ get longer and longer as n grows. Thus Eq. (6) holds.

Theorem 2. *Let $f \colon A^* \to A^*$ be a morphism. If $\mathbf{x} \in A^{\mathbb{N}}$ is a pure morphic word generated by f and $f \in \mathcal{M}_{\mathbf{x}}$, then all extremal words in $\mathcal{S}_{\mathbf{x}}$ are morphic.*

Proof. Let ρ be a total order on A and $b \in A$ a letter occurring in $\mathbf{x}$. We will show that $\mathbf{l}_{b,\rho,\mathbf{x}}$ is morphic. Lemma 9 says that there exist a finite word $w \in A^+$, an infinite word $\mathbf{t} \in \mathcal{S}_{\mathbf{x}}$, and an integer $m \geq 1$ such that $\mathbf{l}_{b,\rho,\mathbf{x}} = w\mathbf{t}$ and either $\mathbf{t} = f^m(\mathbf{t})$ or $\mathbf{t} = \lim_{n\to\infty} x f^m(x) f^{2m}(x) \cdots f^{nm}(x)$ for some finite word $x \in A^+$. Since f^m generates $\mathbf{x}$, the claim that $\mathbf{t}$ is morphic follows in the former case from Theorem 1 and in the latter case from Lemma 5.

Theorem 3. *Let $f\colon A^* \to A^*$ and $g\colon A^* \to B^*$ be morphisms and $\mathbf{x} \in A^{\mathbb{N}}$ such that $f, g \in \mathcal{M}_{\mathbf{x}}$. If $\mathbf{x}$ is a pure morphic word generated by f, then all extremal words in $\mathcal{S}_{g(\mathbf{x})}$ are morphic.*

Proof. Let ρ be a total order ρ on B and $b \in B$. According to Lemma 7, there exists a total order σ on A, a letter $a \in A$, and a word $v \in B^*$ such that

$$\mathbf{s}_{b,\rho,g(\mathbf{x})} = vg(\mathbf{s}_{a,\sigma,\mathbf{x}})$$

Thus it follows from Theorem 2 and Lemma 1 that $\mathbf{l}_{b,\rho,g(\mathbf{x})}$ is morphic.

4 Extremal words of binary pure morphic words

In this section we show that the extremal words of binary pure morphic words are morphic.

Lemma 10. *Let $f\colon \{\mathtt{0},\mathtt{1}\}^* \to \{\mathtt{0},\mathtt{1}\}^*$ be a morphism such that $f(\mathtt{01}) \neq f(\mathtt{10})$. Then $f \in \mathcal{M}_{\mathbf{x}}$ for every $\mathbf{x} \in \{\mathtt{0},\mathtt{1}\}^{\mathbb{N}}$.*

Proof. Let us denote $u = f(\mathtt{0})$ and $v = f(\mathtt{1})$. We have two possibilities:

Case 1. The word u is not a prefix of v^ω. Then there exists an integer $n \geq 0$ such that $u = v^n pas$ and $v = pbt$, where $p, s, t \in \{\mathtt{0},\mathtt{1}\}^*$ and $a, b \in \{\mathtt{0},\mathtt{1}\}$ with $a \neq b$. Now it is easy to see that, for every $\mathbf{y} \in \{\mathtt{0},\mathtt{1}\}^{\mathbb{N}}$, the word $f(\mathtt{1}\mathbf{y})$ begins with $v^n pb$ and $f(\mathtt{0}\mathbf{y})$ begins with $v^n pa$. Therefore $f \in \mathcal{M}_{\mathbf{x}}$ because we may choose $p_{\mathtt{1}} = v^n pb$ and $p_{\mathtt{0}} = v^n pa$.

Case 2. The word u is a prefix of v^ω. Then $v = xy$ and $u = v^n x$ for some integer $n \geq 0$ and words x, y. Now it is easy to see that, for every $\mathbf{y} \in \{\mathtt{0},\mathtt{1}\}^{\mathbb{N}}$, the word $f(\mathtt{0}\mathbf{y})$ begins with $(xy)^n xxy$ and $f(\mathtt{1}\mathbf{y})$ begins with $(xy)^n xyx$. Since $f(\mathtt{01}) \neq f(\mathtt{10})$, it follows that $xy \neq yx$. Denote $xy = pas$ and $yx = pbt$ with a, b distinct letters. Then $f \in \mathcal{M}_{\mathbf{x}}$ because we may let $p_{\mathtt{0}} = (xy)^n xpa$ and $p_{\mathtt{1}} = (xy)^n xpb$.

Theorem 4. *If $\mathbf{x} \in \{\mathtt{0},\mathtt{1}\}^{\mathbb{N}}$ is a binary pure morphic sequence, then all extremal words of $\mathbf{x}$ are morphic.*

Proof. Let f be a binary morphism that generates $\mathbf{x}$. If $f(\mathtt{01}) = f(\mathtt{10})$, then $\mathbf{x}$ is purely periodic, and the claim holds. If $f(\mathtt{01}) \neq f(\mathtt{10})$, then $f \in \mathcal{M}_{\mathbf{x}}$ by Lemma 10, so that $\mathbf{x}$ is morphic by Theorem 2.

There are exactly two total orders on the binary alphabet $\{\mathtt{0},\mathtt{1}\}$; let ρ denote the natural order $\mathtt{0} <_\rho \mathtt{1}$ and $\overline{\rho}$ the other order $\mathtt{1} <_{\overline{\rho}} \mathtt{0}$. The following lemma simplifies the search for the extremal words of a binary pure morphic word, and we will use it later.

Lemma 11. *If $\mathbf{x} \in \{\mathtt{0},\mathtt{1}\}^{\mathbb{N}}$ is a recurrent word in which both $\mathtt{0}$ and $\mathtt{1}$ occur, then*

$$\mathbf{l}_{\mathtt{1},\rho,\mathbf{x}} = \mathtt{1}\mathbf{l}_{\mathtt{0},\rho,\mathbf{x}} \qquad\qquad \mathbf{l}_{\mathtt{0},\overline{\rho},\mathbf{x}} = \mathtt{0}\mathbf{l}_{\mathtt{1},\overline{\rho},\mathbf{x}}. \tag{7}$$

Therefore also

$$\mathbf{s}_{1,\rho,\mathbf{x}} = 0\mathbf{s}_{0,\rho,\mathbf{x}} \qquad\qquad \mathbf{s}_{0,\overline{\rho},\mathbf{x}} = 1\mathbf{s}_{1,\overline{\rho},\mathbf{x}}. \tag{8}$$

Proof. Consider the first equation in (7). On the one hand, $1\mathbf{l}_{0,\rho,\mathbf{x}}$ is in $\mathcal{S}_{\mathbf{x}}$ because the recurrence of $\mathbf{x}$ implies that $a\mathbf{l}_{0,\rho,\mathbf{x}}$ is in $\mathcal{S}_{\mathbf{x}}$ for some $a \in \{0, 1\}$ and if a equaled 0, then the inequality $0\mathbf{l}_{0,\rho,\mathbf{x}} < \mathbf{l}_{0,\rho,\mathbf{x}}$ would contradict the definition of $\mathbf{l}_{0,\rho,\mathbf{x}}$. On the other hand, $1\mathbf{l}_{0,\rho,\mathbf{x}}$ must equal $\mathbf{l}_{1,\rho,\mathbf{x}}$ because otherwise $\mathbf{l}_{1,\rho,\mathbf{x}} < 1\mathbf{l}_{0,\rho,\mathbf{x}}$, which implies $1^{-1}\mathbf{l}_{1,\rho,\mathbf{x}} < \mathbf{l}_{0,\rho,\mathbf{x}}$, and this contradicts the definition of $\mathbf{l}_{0,\rho,\mathbf{x}}$. The second equation in (7) is proved similarly. The identities (8) follow immediately from (7).

5 Extremal words of the Period-doubling word

Let f denote the morphism $0 \mapsto 01$, $1 \mapsto 00$ and let $\mathbf{d} = f^{\omega}(0)$ denote the *period-doubling word* [8,5]. According to Lemma 10, we have $f \in \mathcal{M}_{\mathbf{d}}$.

Let ρ denote the natural order $0 <_{\rho} 1$ and $\overline{\rho}$ the reversed order $1 <_{\overline{\rho}} 0$. Using the observation that neither 0000 nor 11 occur in $\mathbf{d}$ and Lemma 11, the reader has no trouble verifying that the following words start as shown.

$$\mathbf{s}_{0,\rho,\mathbf{d}} = 00100\ldots \qquad\qquad \mathbf{s}_{1,\overline{\rho},\mathbf{d}} = 010100\ldots \tag{9}$$

$$\mathbf{s}_{1,\rho,\mathbf{d}} = 0001\ldots \qquad\qquad \mathbf{s}_{0,\overline{\rho},\mathbf{d}} = 1010100\ldots. \tag{10}$$

Lemma 7 implies that $\mathbf{s}_{0,\rho,\mathbf{d}} = vf(\mathbf{s}_{a,\sigma,\mathbf{d}})$ for some $a \in \{0, 1\}$, proper suffix v of $f(a)$, and $\sigma \in \{\rho, \overline{\rho}\}$. The only possible such factorization has to be of the form $\mathbf{s}_{0,\rho,\mathbf{d}} = 0f(01\ldots)$, so from (9) and (10) we see that $\mathbf{s}_{a,\sigma,\mathbf{d}} = \mathbf{s}_{1,\overline{\rho},\mathbf{d}}$. Thus

$$\mathbf{s}_{0,\rho,\mathbf{d}} = 0f(\mathbf{s}_{1,\overline{\rho},\mathbf{d}}).$$

We can deduce similarly that

$$\mathbf{s}_{1,\overline{\rho},\mathbf{d}} = f(001\ldots) = f(\mathbf{s}_{0,\rho,\mathbf{d}}). \tag{11}$$

Therefore $\mathbf{s}_{0,\rho,\mathbf{d}} = 0f^2(\mathbf{s}_{0,\rho,\mathbf{d}})$, which implies

$$f^2(\mathbf{l}_{0,\rho,\mathbf{d}}) = 01\mathbf{l}_{0,\rho,\mathbf{d}}. \tag{12}$$

We claim that $\mathbf{l}_{0,\rho,\mathbf{d}}$ is the fixed point of the morphism $g\colon 0 \mapsto 0001$ and $1 \mapsto 0101$. Let us denote the unique fixed point of g by $\mathbf{z}$, that is $\mathbf{z} = g^{\omega}(0)$. An easy induction proof shows that $01g(w) = f^2(w)01$ for all $w \in \{0, 1\}^*$. Therefore

$$01\mathbf{z} = 01g(\mathbf{z}) = f^2(\mathbf{z}).$$

Thus by (12), both $\mathbf{z}$ and $\mathbf{l}_{0,\rho,\mathbf{d}}$ satisfy the same relation $01\mathbf{x} = f^2(\mathbf{x})$, which is easily seen to admit a unique solution; thus $\mathbf{z} = \mathbf{l}_{0,\rho,\mathbf{d}}$. Hence, using (11) and Lemma 11, the following result is obtained.

Theorem 5. *Let $\mathbf{d}$ denote the period-doubling word and let $\mathbf{z}$ denote the unique fixed point of the morphism $0 \mapsto 0001$, $1 \mapsto 0101$. Then we have*

$$\mathbf{l}_{0,\rho,\mathbf{d}} = \mathbf{z} \qquad\qquad \mathbf{l}_{1,\rho,\mathbf{d}} = 1\mathbf{z}$$

$$\mathbf{l}_{1,\overline{\rho},\mathbf{d}} = 0^{-1}f(\mathbf{z}) \qquad\qquad \mathbf{l}_{0,\overline{\rho},\mathbf{d}} = f(\mathbf{z}).$$

6 Extremal words of the Chacon word

The Chacon word [9,15] is the fixed point $\mathbf{c} = f^{\omega}(\mathtt{0})$, where f is the morphism $\mathtt{0} \mapsto \mathtt{0010}$, $\mathtt{1} \mapsto \mathtt{1}$. Lemma 10 guarantees that $f \in \mathcal{M}_{\mathbf{x}}$. Let ρ denote the natural order $\mathtt{0} <_{\rho} \mathtt{1}$ and $\overline{\rho}$ the reversed order $\mathtt{1} <_{\overline{\rho}} \mathtt{0}$ as before.

As in Section 5, we use the observation that neither $\mathtt{0000}$ nor $\mathtt{11}$ occur in $\mathbf{c}$ and Lemma 11, to deduce that the following words start as shown.

$$\mathbf{s}_{\mathtt{0},\rho,\mathbf{c}} = \mathtt{001000101}\ldots \qquad \mathbf{s}_{\mathtt{1},\overline{\rho},\mathbf{c}} = \mathtt{010010}\ldots$$
$$\mathbf{s}_{\mathtt{1},\rho,\mathbf{c}} = \mathtt{0001000101}\ldots \qquad \mathbf{s}_{\mathtt{0},\overline{\rho},\mathbf{c}} = \mathtt{1010010}\ldots.$$

Applying Lemma 7 as in the previous section, we find

$$\mathbf{s}_{\mathtt{0},\rho,\mathbf{c}} = f(\mathtt{001}\ldots) = f(\mathbf{s}_{\mathtt{0},\rho,\mathbf{c}}).$$

Since $\mathbf{s}_{\mathtt{0},\rho,\mathbf{c}}$ begins with $\mathtt{0}$, we thus have $\mathbf{s}_{\mathtt{0},\rho,\mathbf{c}} = \mathbf{c}$ and $\mathbf{l}_{\mathtt{0},\rho,\mathbf{c}} = \mathtt{0}\mathbf{c}$.

Similarly, recalling that $\mathbf{s}_{\mathtt{0},\overline{\rho},\mathbf{c}} = \mathtt{1}\mathbf{s}_{\mathtt{1},\overline{\rho},\mathbf{c}}$ by Lemma 11, we deduce using Lemma 7 that

$$\mathbf{s}_{\mathtt{1},\overline{\rho},\mathbf{c}} = \mathtt{0}f(\mathtt{10}\ldots) = \mathtt{0}f(\mathbf{s}_{\mathtt{0},\overline{\rho},\mathbf{c}}) = \mathtt{01}f(\mathbf{s}_{\mathtt{1},\overline{\rho},\mathbf{c}}).$$

Therefore $\mathbf{l}_{\mathtt{1},\overline{\rho},\mathbf{c}}$ can be expressed as $\mathbf{l}_{\mathtt{1},\overline{\rho},\mathbf{c}} = \tau g^{\omega}(b)$, where b is a new symbol, g is a morphism for which $g(b) = b\mathtt{01}$ and $g(a) = f(a)$ for $a \in \{\mathtt{0}, \mathtt{1}\}$, and $\tau(b) = \mathtt{1}$ and $\tau(a) = a$ for $a \in \{\mathtt{0}, \mathtt{1}\}$. Thus a final application of Lemma 11 allows us to wrap up the results of this section as follows.

Theorem 6. *Let* $\mathbf{c}$ *denote the Chacon word. Then we have*

$$\mathbf{l}_{\mathtt{0},\rho,\mathbf{c}} = \mathtt{0}\mathbf{c} \qquad \mathbf{l}_{\mathtt{1},\overline{\rho},\mathbf{c}} = \tau g^{\omega}(b)$$
$$\mathbf{l}_{\mathtt{1},\rho,\mathbf{c}} = \mathtt{10}\mathbf{c} \qquad \mathbf{l}_{\mathtt{0},\overline{\rho},\mathbf{c}} = \mathtt{0}\tau g^{\omega}(b),$$

where g *and* τ *are the morphisms given above.*

7 The least word in the shift orbit closure of the Rudin-Shapiro word

In this section, we give a new proof for the form of the lexicographically smallest word in the shift orbit closure of the Rudin-Shapiro word. This result was first derived in [7]. Considerations in this section are more involved than the ones in the previous sections because a coding is needed in the definition of the Rudin-Shapiro word. In what follows, we denote the natural order on letters $\mathtt{0}, \mathtt{1}, \mathtt{2}, \mathtt{3}$ by ρ. Thus we have $\mathtt{0} <_{\rho} \mathtt{1} <_{\rho} \mathtt{2} <_{\rho} \mathtt{3}$.

Let f and g be the morphisms

$$f\colon \begin{cases} \mathtt{0} \mapsto \mathtt{01} \\ \mathtt{1} \mapsto \mathtt{02} \\ \mathtt{2} \mapsto \mathtt{31} \\ \mathtt{3} \mapsto \mathtt{32} \end{cases} \quad \text{and} \quad g\colon \begin{cases} \mathtt{0} \mapsto \mathtt{0} \\ \mathtt{1} \mapsto \mathtt{0} \\ \mathtt{2} \mapsto \mathtt{1} \\ \mathtt{3} \mapsto \mathtt{1} \end{cases}$$

Denote

$$\mathbf{u} = f^\omega(\mathtt{0}) = \mathtt{0102013101023202010201313231013101020131}\ldots$$

and

$$\mathbf{w} = g(\mathbf{u}) = \mathtt{0001001000011101000100101110001000010010}\ldots.$$

Then $\mathbf{w}$ is the Rudin-Shapiro word, and our goal is to prove the identity $\mathbf{l}_{\mathtt{0},\rho,\mathbf{w}} = \mathtt{0}\mathbf{w}$. To that end, we need the next two lemmas. Let us denote $\Sigma_4 = \{\mathtt{0},\mathtt{1},\mathtt{2},\mathtt{3}\}$.

Lemma 12. *Let σ and σ' be two total orders on Σ_4. If σ and σ' order the pairs $(\mathtt{0},\mathtt{3})$ and $(\mathtt{1},\mathtt{2})$ in the same way, i.e., $\mathtt{0} <_\sigma \mathtt{3}$ iff $\mathtt{0} <_{\sigma'} \mathtt{3}$ and $\mathtt{1} <_\sigma \mathtt{2}$ iff $\mathtt{1} <_{\sigma'} \mathtt{2}$, then $\mathbf{l}_{d,\sigma,\mathbf{u}} = \mathbf{l}_{d,\sigma',\mathbf{u}}$ for all $d \in \Sigma_4$.*

Proof. Suppose that $\mathbf{l}_{d,\sigma,\mathbf{u}} = ua\mathbf{t}$ and $\mathbf{l}_{d,\sigma',\mathbf{u}} = ub\mathbf{t}'$ with distinct letters $a, b \in \Sigma_4$. Since σ and σ' agree on $(\mathtt{0},\mathtt{3})$ and $(\mathtt{1},\mathtt{2})$, it follows that either $a \in \{\mathtt{0},\mathtt{3}\}$ and $b \in \{\mathtt{1},\mathtt{2}\}$, or vice versa. Furthermore, if c denotes the last letter of u, then both ca and cb occur in $\mathbf{u}$. This contradicts the fact that none of the words $\mathtt{00}$, $\mathtt{03}$, $\mathtt{11}$, $\mathtt{12}$, $\mathtt{21}$, $\mathtt{22}$, $\mathtt{30}$, $\mathtt{33}$ occur in $\mathbf{u}$.

The next lemma is interesting in its own right. It was also proved in [7].

Lemma 13. *We have $\mathbf{l}_{\mathtt{0},\rho,\mathbf{u}} = \mathbf{u}$.*

Proof. Since clearly $f \in \mathcal{M}_\mathbf{u}$, Lemma 7 implies that there exist a letter $a \in \Sigma_4$, a proper suffix v of $f(a)$, and a total order σ on Σ_4 such that $\mathbf{s}_{\mathtt{0},\rho,\mathbf{u}} = vf(\mathbf{s}_{a,\sigma,\mathbf{u}})$. An easy case analysis based on the observation that $\mathtt{00}$ does not occur in $\mathbf{u}$ yields $\mathbf{s}_{\mathtt{0},\rho,\mathbf{u}} = \mathtt{10201}\ldots = \mathtt{1}f(\mathtt{10}\ldots)$, and hence

$$v = \mathtt{1} \qquad \text{and} \qquad \mathbf{s}_{a,\sigma,\mathbf{u}} = \mathtt{10}\ldots.$$

Since $v = \mathtt{1}$ is a suffix of $f(a)$, we have $a = \mathtt{0}$ or $a = \mathtt{2}$. Furthermore since $\mathbf{l}_{a,\sigma,\mathbf{u}}$ starts with $a\mathtt{1}$, we must have $a = \mathtt{0}$ because $\mathtt{21}$ does not occur in $\mathbf{u}$. Thus $\mathbf{s}_{\mathtt{0},\rho,\mathbf{u}} = \mathtt{1}f(\mathbf{s}_{\mathtt{0},\sigma,\mathbf{u}})$.

Next we claim $\mathbf{s}_{\mathtt{0},\sigma,\mathbf{u}} = \mathbf{s}_{\mathtt{0},\rho,\mathbf{u}}$. We prove this by showing that $\mathtt{0} <_\sigma \mathtt{3}$ and $\mathtt{1} <_\sigma \mathtt{2}$; then the claim follows from Lemma 12. If, contrary to what we want to show, we have $\mathtt{2} <_\sigma \mathtt{1}$, then $\mathbf{l}_{\mathtt{0},\sigma,\mathbf{u}}$ would begin with $\mathtt{02}$, contradicting the fact that $\mathbf{s}_{\mathtt{0},\sigma,\mathbf{u}}$ begins with $\mathtt{1}$. Consequently we have $\mathtt{1} <_\sigma \mathtt{2}$. Furthermore if $\mathtt{3} <_\sigma \mathtt{0}$, then $\mathbf{l}_{\mathtt{0},\sigma,\mathbf{u}}$ would begin with $\mathtt{013}$, contradicting the fact that $\mathbf{s}_{\mathtt{0},\sigma,\mathbf{u}}$ begins with $\mathtt{10}$. Therefore $\mathbf{s}_{\mathtt{0},\sigma,\mathbf{u}} = \mathbf{s}_{\mathtt{0},\rho,\mathbf{u}}$.

Now the identity $\mathbf{s}_{\mathtt{0},\rho,\mathbf{u}} = \mathtt{1}f(\mathbf{s}_{\mathtt{0},\rho,\mathbf{u}})$ implies $\mathbf{l}_{\mathtt{0},\rho,\mathbf{u}} = f(\mathbf{l}_{\mathtt{0},\rho,\mathbf{u}})$, so that $\mathbf{l}_{\mathtt{0},\rho,\mathbf{u}}$ is the unique iterative fixed point of f that starts with $\mathtt{0}$, that is $\mathbf{l}_{\mathtt{0},\rho,\mathbf{u}} = \mathbf{u}$.

Finally, we are ready to prove the main result of this subsection.

Theorem 7. *Let $\mathbf{w}$ denote the Rudin–Shapiro word. Then $\mathbf{l}_{\mathtt{0},\rho,\mathbf{w}} = \mathtt{0}\mathbf{w}$.*

Proof. Let $h = g \circ f$ be the composition of g and f. Then

$$h\colon \quad \mathtt{0} \mapsto \mathtt{00} \qquad \mathtt{1} \mapsto \mathtt{01} \qquad \mathtt{2} \mapsto \mathtt{10} \qquad \mathtt{3} \mapsto \mathtt{11}.$$

According to Lemma 6, there exist a letter $a \in \Sigma_4$ and an infinite word $\mathbf{z} \in \Sigma_4^{\mathbb{N}}$ such that $a\mathbf{z} \in \mathcal{S}_{\mathbf{u}}$ and $\mathbf{l}_{\mathtt{0},\rho,\mathbf{w}} = uh(\mathbf{z})$, where u is a nonempty suffix of $h(a)$. Since $\mathbf{l}_{\mathtt{0},\rho,\mathbf{w}}$ clearly starts with $\mathtt{0000}$ and $\mathtt{00}$ does not occur in $\mathbf{u}$, it follows that $u = \mathtt{0}$, $\mathbf{z} = \mathtt{01}\ldots$, and $a = \mathtt{2}$.

On the other hand, it is easy to see that $\mathtt{2}\mathbf{u} \in \mathcal{S}_{\mathbf{u}}$. Since $\mathbf{u} = \mathbf{l}_{\mathtt{0},\rho,\mathbf{u}}$ by Lemma 13, we have $\mathbf{u} \leq_\rho \mathbf{z}$, and so $\mathtt{2}\mathbf{u} \leq_\rho \mathtt{2}\mathbf{z}$. Furthermore, since h preserves ρ, that is to say if $x, y \in \{\mathtt{0}, \mathtt{1}, \mathtt{2}, \mathtt{3}\}^*$ with $x <_\rho y$, then $h(x) <_\rho h(y)$, we have $h(\mathtt{2}\mathbf{u}) \leq_\rho h(\mathtt{2}\mathbf{z})$, which gives $\mathtt{0}h(\mathbf{u}) \leq_\rho \mathtt{0}h(\mathbf{z}) = \mathbf{l}_{\mathtt{0},\rho,\mathbf{w}}$. Hence we must have $\mathtt{0}h(\mathbf{u}) = \mathbf{l}_{\mathtt{0},\rho,\mathbf{w}}$, and so

$$\mathbf{l}_{\mathtt{0},\rho,\mathbf{w}} = \mathtt{0}h(\mathbf{u}) = \mathtt{0}g(\mathbf{u}) = \mathtt{0}\mathbf{w}.$$

8 Conclusion

The condition $f, g \in \mathcal{M}_{\mathbf{x}}$ guaranteeing that the extremal words of a morphic word of the form $g\big(f^\omega(a)\big)$ are morphic (Theorem 3) is quite powerful, as we have seen. Clearly, however, not all morphic sequences and the corresponding morphisms satisfy this condition. So does there exist a morphic sequence with an extremal sequence that is not morphic? All our failed attempts to produce such an example encourage us to conjecture that, in fact, all extremal words of all morphic sequences are morphic.

References

1. J.–P. Allouche. Théorie des nombres et automates, Thèse d'État, Université Bordeaux I, 1983.
2. J.–P. Allouche, J. Currie, and J. Shallit. Extremal infinite overlap-free binary words. *The Electronic Journal of Combinatorics* **5** (1998), #R27.
3. J.–P. Allouche and M. Cosnard. Itérations de fonctions unimodales et suites engendrées par automates, *C. R. Acad. Sci. Paris, Sér. A* **296** (1983), 159–162.
4. J.–P. Allouche, N. Rampersad, and J. Shallit. Periodicity, repetitions, and orbits of an automatic sequence. *Theoret. Comput. Sci.* **410** (2009), 2795–2803.
5. J.–P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations.* Cambridge University Press, 2003.
6. J. Cassaigne and F. Nicolas. *Factor Complexity* in *Combinatorics, Automata and Number Theory*, V. Berthé and M. Rigo (eds.), Cambridge Unicersity Press, Cambridge, 2010.
7. J. Currie. Lexicographically least words in the orbit closure of the Rudin–Shapiro word. *Theoret. Comput. Sci.* **412** (2011), 4742–4746.
8. D. Damanik. Local symmetries in the period-doubling sequence. *Discrete Appl. Math.* **100** (2000), 115–121.
9. S. Ferenczi. Les transformations de Chacon: combinatoire, structure géométrique, lien avec les systèmes, de complexité $2n+1$. *Bulletin de la S. M. F.*, tome 123, n°2 (1995), p. 271– 292.

10. S. Gan. Sturmian sequences and the lexicographic world. *Proc. Amer. Math. Soc.* **129**(5) (2000), 1445–1451 (electronic).
11. T. Head and B. Lando. Fixed and stationary ω-words and ω-languages in *The Book of L*, G. Rozenberg and A. Salomaa (eds.), pp. 147–156, Springer-Verlag, 1986.
12. D. Krieger. On stabilizers of infinite words. *Theoret. Comput. Sci.* **400** (2008), 169–181.
13. M. Lothaire. *Algebraic Combinatorics on Words* in Encyclopedia of Mathematics and its Applications, vol. 90, Cambride University Press, Cambridge, 2002.
14. G. Pirillo. Inequalities characterizing standard Sturmian and episturmian words. *Theoret. Comput. Sci.* **341** (2005), 276–292.
15. N. Pytheas Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics*, V. Berthé, S. Ferenczi, C. Mauduit, and A. Siegel (eds.), Vol. 1794 of Lecture Notes in Mathematics, Springer, 2002.